REMARKS

Claim Status

Applicants thank the Examiner, Mr. James S. Wozniak, for his courtesies extended to their representative during the telephonic interview conducted on June 19, 2007, and for his assistance in furthering prosecution on the merits of the instant application. During the telephonic interview, proposed amendments to independent claims 1 and 11 were discussed. It was agreed that the proposed amendments to these claims overcame the rejections under 35. U.S.C. §112 contained in the Office Action dated January 11, 2007. However, no agreement with respect to patentability of the claims over the prior art was reached. The following comments expand on the subject matter discussed during the telephonic interview.

Claims 1 and 4-12 are currently pending, with claims 1 and 11 being in independent form. Claims 1 and 11 have been amended. No new matter has been added. Reconsideration of the application, as amended, is respectfully requested.

Overview of the Office Action

Claims 1 and 4-12 have been objected to based on minor informalities. Withdrawal of this objection is in order, as explained below.

Claims 1, 4-10 and 12 stand rejected under 35 U.S.C. §112, first paragraph, as failing to comply with the enablement requirement. Withdrawal of these rejections is in order, as also explained below.

Claims 1 and 4-12 stand rejected under 35 U.S.C. §112, first paragraph, as failing to comply with the written description requirement. Withdrawal of these rejection is in order, as also explained below.

Claims 1 and 4-12 stand rejected under 35 U.S.C. §112, second paragraph, as indefinite for failing to particularly point out and distinctly claim the subject matter which Applicant regards as the invention. Withdrawal of this rejection is also in order, as explained below.

Claims 1 and 9-10 stand rejected under 35 U.S.C. §103(a) as obvious over U.S. Patent No. 5,699,481 ("*Shlomot I*") in view of U.S. Patent No. 5,694,521 ("*Shlomot II*"). Claims 4-6 and 11 stand rejected under 35 U.S.C. §103(a) as obvious over *Shlomot I* in view of *Shlomot II*, and further in view of U.S. Patent No. 5,825,771 ("*Cohen*"). Claim 7 stands rejected under 35 U.S.C. §103(a) as obvious over *Shlomot I* in view of *Shlomot II*, further in view of *Cohen*, and further in view of U.S. Patent No. 5,897,613 ("*Chan*"). Claim 8 stands rejected under 35 U.S.C. §103(a) as obvious over *Shlomot I* in view of *Shlomot II*, and further in view of U.S. Patent No. 5,642,466 ("*Narayan*"). Lastly, claim 12 stands rejected under 35 U.S.C. §103(a) as obvious over *Shlomot I* in view of *Shlomot II*, and further in view of U.S. Patent No. 5,696,875 ("*Pan*").

Applicants have carefully considered the Examiner's rejections and the comments provided in support thereof. For the following reasons, Applicants assert that all claims now presented for examination in the present application are patentable over the cited art.

Amendments Addressing Section 112 Issues and Formalities

The Examiner stated in "claim 1, line 9 and claim 11, line 16, it appears that 'and' should be changed to –or-- because it is unclear whether the two buffer level conditions should be referred to in the alternative or can both occur simultaneously". In response to this objection, Applicants have amended independent claims 1 and 11 to indicate that "implementing voice activity detection" and "implementing concatenation processing" are performed whenever the filling level lies between the particular threshold levels. As described at pg. 23, lines 14-19 of

the originally filed specification, "it is possible to decide whether or not to activate particular alarm levels. For example, when asynchrony is "weak" then alarm levels 1 and 2 can be combined, and the simple solution of replacing two frames by a single frame can then constitute the only active process". In view of the foregoing, claims 1 and 11 are definite, and withdrawal of the objection is therefore in order.

Claims 1 and 4-12 stand rejected under 35 U.S.C. §112, second paragraph. In response to each specific rejection, Applicants have amended claims 1 and 11 in a manner which is believed to be self-explanatory. Accordingly, withdrawal of these rejections is also appropriate.

Patentability of the Claims under 35 U.S.C. §112, First Paragraph

The Examiner (pg. 4, paragraph 7 of the Office Action) stated that "claim 1 recites a single step method for eliminating non-active frames *or* concatenating frames based on a filling but lacks means for enabling those operations". In response to this objection, claim 1 has been amended in a manner that is believed to be self-explanatory. That is, claim 1 has been amended to recite the steps associated with Applicants' inventive method. Withdrawal of this rejection is therefore in order.

The Examiner stated that the claimed limitation "'concatenation processing is implemented irrespective of the content of the frames' is not supported by the specification and fails to comply with the written description". In response to this rejection, Applicants have amended claims 1 and 11 in a manner that is believed to be self-explanatory. Withdrawal of this rejection is therefore in order.

Descriptive Summary of the Prior Art

Shlomot I discloses "a timing recovery scheme for packet speech in communication systems". According to *Shlomot I*, "the timing recovery scheme is performed by the speech decoder and is transparent to the controller operation" (see col. 3, lines 45-49).

Shlomot II discloses "a variable speed playback system incorporating multiple-period template matching to alter the LPC excitation periodical structure, and thereby increase or decrease the rate of speech playback, while retaining the natural quality of the speech" (see col. 1, line 66 thru col. 2, line 3).

Cohen discloses "an audio transceiver (having an audio receiver and a transmitter) which, on the receiving side, adaptively controls the amount of audio data in the buffer of a PC audio device, such that the audio device always has something to play. On the transmission side, the audio transmitter provides at least sequence numbers to the audio packets to be sent. The audio receiver receives the audio packets, processes them and plays them as soon as possible thereafter" (see col. 2, lines 44-53).

Chan discloses "a silence-suppression scheme for a packet voice transmission system [that] accurately regenerates the silence signals at the receiving end" of a transmission (see Abstract).

Narayan discloses "a software-only real time text-to-speech system including intonation control which does not introduce discontinuities into output speech stream" (see col. 2, lines 39-41).

Pan discloses a method and system for compressing a speech signal into compressed speech data (see col. 1, lines 55-57).

Patentability of Independent Claim 1 under 35 U.S.C. §103(a)

Independent claim 1 has been amended to recite the step of "implementing concatenation processing on two successive frames to compact the two frames into a pseudo-frame of length less than or equal to one frame whenever the filling level lies between the second threshold and a third threshold, a length reduction ratio of the pseudo-frame relative to the length of the two frames being greater than or equal to two. Independent claim 11 has been correspondingly amended. No new matter has been added.

The Examiner (pg. 4 of the Office Action) acknowledges that *Shlomot I* fails to teach or suggest that "the further processing comprises concatenation processing used to compact two successive frames into a pseudo-frame having a reduction ratio greater than or equal to two," as recited in independent claim 1, and cites *Shlomot II* for this feature. The combination of *Shlomot I* and *Shlomot II*, however, fails to achieve Applicants' claimed method; there is nothing in *Shlomot II* to cure the above-noted deficiency in *Shlomot I.*

A person having the ordinary level of skill in the art of speech processing would possess the understating that in code excited linear predictive (CELP) coding, a set of Gaussian codes from a code book is used to extract the vocal tract model from the speech signal. Here, a code that minimizes the signal "residue" is selected as a correct excitation function and the corresponding code index is then transmitted. For example, *Shlomot II* (col. 1, lines29-33) states that "LPC techniques may be [used] for speech coding involving code excited linear prediction (CELP) speech coders. These conventional speech coders generally utilize at least two excitation codebooks. The outputs of the codebooks provide the input to the LPC synthesis filter".

Shlomot II (col. 1, lines 51-56) describes alternative ways to achieve typical store and retrieve operations, speech modification, such as fast and slow playback, using a variety of time domain and frequency domain estimation and modification techniques, where several speech parameters are estimated, e.g., pitch frequency or lag, and the speech signal is accordingly modified. *Shlomot II* (col. 1, lines 60-63) teaches that simpler and more robust speech modification can be achieved by utilizing template matching instead of pitch estimation.

Bearing the foregoing in mind, *Shlomot II* (col. 1, lines 66 thru col. 2, line 3) is directed to "a variable speed playback system incorporating multiple-period template matching to alter the LPC excitation periodical structure, and thereby increase or decrease the rate of speech playback, while retaining the natural quality of the speech".

According to *Shlomot II*, "[a] multiple-period similarity measure is determined for a decoded LPC excitation signal" (see col. 2, lines 6-8). *Shlomot II* (col. 2, lines 8-11) describes that "[e]xpansion or compression of the time domain LPC excitation signal may then be performed according to a rational factor, e.g., 1:2, 2:3, 3:4, 4:3, 3:2, and 2:1". *Shlomot II* (col. 2, lines 11-14) further describes that "[t]he expansion and compression are performed on the LPC excitation signal...". *Shlomot II* thus clearly teaches that the excitation signal is compressed.

Shlomot II (col. 2, lines 17-27) teaches that at least two templates of the LPC excitation signal are determined, the templates are defined by one or more segments within the LPC excitation signal, two complementary windows are constructed, these templates are then multiplied by the windows, overlapped and summed, and the resultant excitation signal representing modified excitation signal is input into an LPC synthesis filter. Clearly, *Shlomot II* teaches that it is the excitation signal applied to an LPC filter that is processed or compressed.

However, there is no concatenation of two successive frames into a pseudo-frame in the manner required by Applicants' amended independent claim 1.

The Examiner (pg. 8 of the Office Action) asserts that:

> Shlomot II ... recites a fast speech data playback method that compresses two consecutive speech segments into a single segment, irrespective of the segment content, by a ratio of greater than or equal to 2 *(Col. 4, Line 40- Col. 5, Line 14).*

Applicants respectfully disagree with the Examiner's assertion. *Shlomot II* (col. 3, lines 4-5; FIGS. 3(a) thru 3(c)) teaches that "Tstart marks the beginning of [an] X_{ML} template". *Shlomot II* (col. 3, lines 5-10) further describes that "[a]t each stage, properly aligned templates X_{ML} and y_{ML} of [an] excitation signal **302** are correlated ... for each possible integer value L between a minimum number Lmin to a maximum Lmax", where a normalized correlation is represented by an equation, i.e., Eqn. 1. *Shlomot II* (col. 3, lines 41-43) states that "L^* represents the periodical structure of the excitation signal, and in most cases coincides with the pitch period". After the performance of additional processing steps, *Shlomot II* (col. 3, lines 41-43) states that "the excitation signal can then be filtered by the LPC synthesis filter **104** (FIG. 1) to produce the decoded output speech **108**".

Throughout the entire foregoing text, *Shlomot II* describes the processing of an excitation signal. However, there is no specific teaching or suggestion that two successive frames are concatenated together. *Shlomot II* makes no mention of such a concept whatsoever. In fact, col. 3, lines 24-27 of *Shlomot II* is the only context in which the word frame even appears, where it is stated that "normalized correlation is not confined to the usual frame structure used in LPC/CELP coding, and L^* is not necessarily limited to the pitch period". However, this provides nothing with respect to how the "usual frame structure used in LPC/CELP" is processed, apart from the compression of an excitation signal to which *Shlomot II* is directed. The skilled person

understands that a frame is data that is transmitted between specific communication points as a unit complete with addressing and necessary protocol control information. Generally, a frame is transmitted serially bit-by-bit and contains a header field and a trailer field that "frame" the data. For example, a flag and address fields may constitute the header of a frame. A frame check sequence and second flag fields may constitute a trailer of the frame. The information or data in the frame may contain another encapsulated frame that is used in a higher-level or different protocol, for example. *Shlomot II* therefore does not teach frame manipulation.

Shlomot II (col. 4, line 40-61) specifically describes how "*Fast Playback--Excitation Signal Compression*" is achieved (Emphasis Added). *Shlomot II* (col. 4, line 41-43; FIG. 3(a)) teaches "data compression at a 2-to-1 ratio ... is achieved by combining the templates x_L and y_L into one template of length L". According to *Shlomot II*, "[t]emplate x_L **312** is defined by the L samples starting from Tstart, and y_L is defined by the next segment of L samples". However, it is important to note that Applicants claim the concatenation of two successive frames, as opposed to combining templates of length L.

Shlomot II (col. 4, lines 46-51) further describes that "[f]or each L in the range Lmin to Lmax, the normalized correlation C_L, is calculated according to Eqn. (1), where M=1, and L^* is chosen as the value of L which maximizes the normalized correlation". As further described in *Shlomot II*, "x_{L^*} is multiplied by W_{L^*} (**402**) and y_{L^*} is multiplied by $W^Y_{L^*}$ (**404**). The resulting signals are then overlapped (**406**) and summed (**408**), yielding the compressed excitation signal (**410**)" (see col. 4, lines 52-56; FIG. 4). Thus, according to *Shlomot II*, since two non-overlapped segments of L^* samples each are combined into one segment of L^* samples, 2-to-1 compression is achieved (see col. 4, lines 56-58; FIG. 3(a)). However, the 2-1 compression of an excitation signal is not the concatenation of two successive frames, as recited and claimed in amended

independent claim 1. Clearly, *Shlomot II* teaches compression of an excitation signal which clearly differs from what is disclosed and claimed by Applicants. Therefore, amended independent claim 1 is patentable over the combination of *Shlomot I and Shlomot II* for at least this reason.

In addition, absent an impermissible hindsight analysis based on Applicants disclosure, the skilled person would not seek to modify the method and system taught in *Shlomot I* based on the teachings of *Shlomot II*. The claimed invention is directed to a method for managing decoding and playback of a sound signal. Indeed, amended claim 1 encompasses any kind of sound signal, including music signals (see pg. 1, line 21 and pg. 25, lines 26-29 of the originally filed specification). *Shlomot I*, however, is solely directed to processing speech signals.

Shlomot I discloses a method for handling bit stream frames that contain speech. *Shlomot I* takes advantage of the existence of silence during speech to provide for corrections when too many or too few frames are received in a given time period. When too many frames are received, frames containing silence are deleted and when too few are received, frames containing silence are repeated.

As defined by amended independent claim 1, any overabundance of a filling level of at least one of a first and a second buffer memory is detected by comparing the filling level with at least one threshold such that if the filling level lies between a second threshold and a third threshold, concatenation processing is implemented on two successive frames to compact them into a pseudo-frame of length less than or equal to one frame, where the length reduction ratio of the pseudo-frame relative to the length of the two frames is greater than or equal to two.

As noted previously and conceded by the Examiner, *Shlomot I* fails to teach or suggest implementing concatenation processing on two successive frames to compact the two frames into

a pseudo-frame of length less than or equal to one frame. Consequently, the problem to be solved by the skilled person looking to modify *Shlomot I* to achieve the result obtained by Applicants, which *Shlomot I* fails to achieve, is how to produce a continuous stream of a decoded sound (i.e., any kind of sound) in an asynchronous transmission system.

Shlomot II discloses a variable speed playback method for increasing or decreasing the rate of speech that is played back. More particularly, two templates of length L are identified in an LPC excitation signal and a data compression ratio of at least 2-to-1 is achieved by combining the two templates into one template of length L. *Shlomot II* fails to teach or suggest a method of managing the decoding and playback of a sound signal in an asynchronous transmission system as encompassed by amended claim 1. There is simply no incentive, teaching or reason that would provide the skilled person with the motivation to modify or adapt the method of *Shlomot I* to a method falling within the terms of the claimed solution. In fact, without considerable effort, *Shlomot II* cannot readily be combined with *Shlomot I* because the systems disclosed in each of these patent are completely incompatible.

Shlomot II teaches the manipulation of the periodical structure of speech signals to improve digital storage and retrieval systems. In particular, *Shlomot II* is directed to processing voice by compressing an excitation signal to allow electronic devices to store and playback digital messages. *Shlomot II*, however, has nothing to do with a decoded sound in an asynchronous transmission system. Thus, it would not have been obvious to a person of ordinary skill in the art at the time of the invention to look to *Shlomot II* in order to solve the above described problem that *Shlomot I* fails to solve. Therefore, the combination of *Shlomot I* and *Shlomot II* is improper. In view of the foregoing, reconsideration and withdrawal of the rejection

under 35 U.S.C. §103(a) are respectfully requested, and a notice to that effect is earnestly solicited.

Patentability of Independent Claim 11 under 35 U.S.C. §103(a)

The foregoing comments relating to the failure of *Shlomot I* and *Shlomot II* to teach or suggest the concatenation of two successive frames into a pseudo-frame are equally pertinent to demonstrating the patentability of independent claim 11.

The Examiner (pg. 10 of the Office Action) acknowledges that the combination of *Shlomot I* and *Shlomot II* fails to teach or suggest "the use of a playback buffer," as recited in independent claim 11, and cites *Cohen* for this feature. The combination of *Shlomot I*, *Shlomot II* and *Cohen*, however, fails to achieve Applicants' claimed system; there is nothing in *Cohen* to cure the above-noted deficiency in *Shlomot I* and *Shlomot II*. That is, *Cohen* fails to teach or suggest "concatenation of two successive frames into a pseudo-frame" in the manner required by Applicants' amended independent claim 11.

In view of the foregoing, independent claim 11 is patentable over the combination of *Shlomot I*, *Shlomot II* and *Cohen* on at least this basis. Withdrawal of the rejection under 35 U.S.C. §103(a) is therefore requested, and a notice to that effect is earnestly solicited.

Patentability of the Dependent Claims over the Prior Art under 35 U.S.C. 103

The Examiner cites *Cohen* in an attempt to cure the shortcomings of the combination of *Shlomot I* and *Shlomot II*, i.e., "decoding processing in a cyclical manner relative to the content of the first buffer memory" and "a fake frame is not generated when a missing or erroneous frame is detected for a frame on which an absence of samples has already been detected" as

recited in dependent claims 5 and 6, respectively. The Examiner cites *Chan* in an attempt to cure the shortcomings of the combination of *Shlomot I*, *Shlomot II* and *Cohen*, i.e., "a previously stored frame to determine the generation of a correction frame," as recited in dependent claim 7. The Examiner cites *Narayan* in an attempt to cure the shortcomings of the combination of *Shlomot I* and *Shlomot II*, i.e., "the weighting scheme for [combining] speech segments," as recited in dependent claim 8. Lastly, the Examiner cites *Pan* in an attempt to cure the shortcomings of the combination of *Shlomot I* and *Shlomot II*, i.e., "averaging combined speech segments," as recited in dependent claim 12. However, each of these references fails to cure the deficiency of the method achieved by the combination of *Shlomot I* and *Shlomot II*, because *Cohen*, *Chan*, *Narayan* and/or *Pan*, individually or in combination, fail to teach or suggest concatenation of two successive frames into a pseudo-frame as recited in amended independent claim 1. Therefore, dependent claims 5-8 and 12 are patentable.

In view of the patentability of independent claims 1 and 11, for the reasons presented above, each of dependent claims 4-10 and 12 is patentable therewith. Moreover, each of these claims includes features which serve to even more clearly distinguish the invention over the applied references.

Conclusion

Based on all of the above, it is respectfully submitted that the present application is now in proper condition for allowance. Prompt and favorable action to this effect and early passing of this application to issue are respectfully solicited.

Should the Examiner have any comments, questions, suggestions or objections, the Examiner is respectfully requested to telephone the undersigned in order to facilitate reaching a resolution of any outstanding issues.

Respectfully submitted,
COHEN PONTANI LIEBERMAN & PAVANE LLP

By 

Edward M. Weisz
Reg. No. 37,257
551 Fifth Avenue, Suite 1210
New York, New York 10176
(212) 687-2770

Dated: June 19, 2007